

0570
0103

#5



OIEP

RAW SEQUENCE LISTING

DATE: 01/09/2003

PATENT APPLICATION: US/10/023,649

TIME: 13:34:41

Input Set : D:\NEL-0001 Sequence Listing.txt

Output Set: N:\CRF4\01092003\J023649.raw

3 <110> APPLICANT: The Minister of National Defence, Government of Canada
 4 Nagata, Leslie P
 5 Wong, Jonathon P
 7 <120> TITLE OF INVENTION: Novel DNA-Based Vaccine Against the Encephalitis
 Alphaviruses
 9 <130> FILE REFERENCE: NEL-001
 11 <140> CURRENT APPLICATION NUMBER: 10/023,649
 C--> 12 <141> CURRENT FILING DATE: 2002-12-27
 14 <150> PRIOR APPLICATION NUMBER: 60/256,948
 15 <151> PRIOR FILING DATE: 2000-12-21
 17 <160> NUMBER OF SEQ ID NOS: 7
 19 <170> SOFTWARE: PatentIn version 3.1
 21 <210> SEQ ID NO: 1
 22 <211> LENGTH: 11484
 23 <212> TYPE: DNA
 24 <213> ORGANISM: Western equine encephalomyelitis virus - strain 71V-1658
 26 <220> FEATURE:
 27 <221> NAME/KEY: CDS
 28 <222> LOCATION: (25)..(7428)
 29 <223> OTHER INFORMATION: 5' UTR <1 .. 24
 32 <220> FEATURE:
 33 <221> NAME/KEY: CDS
 34 <222> LOCATION: (7473)..(11183)
 35 <223> OTHER INFORMATION: encodes nucleocapsid, E3, E2, 6K and E1 proteins
 38 <400> SEQUENCE: 1
 39 accctacaaa ctaatcgatc caat atg gaa aga att cac gtt gac tta gat 51
 40 Met Glu Arg Ile His Val Asp Leu Asp
 41 1 5
 43 gct gac agc ccg tat gtc aag tcg tta cag cgg acg ttt cca caa ttt 99
 44 Ala Asp Ser Pro Tyr Val Lys Ser Leu Gln Arg Thr Phe Pro Gln Phe
 45 10 15 20 25
 47 gag atc gaa gca agg cag gtc act gac aat gac cat gcc aat gcc aga 147
 48 Glu Ile Glu Ala Arg Gln Val Thr Asp Asn Asp His Ala Asn Ala Arg
 49 30 35 40
 51 gcg ttt tcg cat gtg gca aca aag ctc att gag agc gaa gtc gac cgg 195
 52 Ala Phe Ser His Val Ala Thr Lys Leu Ile Glu Ser Glu Val Asp Arg
 53 45 50 55
 55 gac caa gtt atc ttg gac att gga agt gcg ccc gtc aga cat gca cat 243
 56 Asp Gln Val Ile Leu Asp Ile Gly Ser Ala Pro Val Arg His Ala His
 57 60 65 70
 59 tcc aat cac cgc tat cat tgt atc tgc cct atg ata agc gct gaa gac 291
 60 Ser Asn His Arg Tyr His Cys Ile Cys Pro Met Ile Ser Ala Glu Asp
 61 75 80 85
 63 ccg gac aga cta caa cgg tat gca gaa aga ctt aag aaa agt gac att 339

ENTERED

RAW SEQUENCE LISTING

DATE: 01/09/2003

PATENT APPLICATION: US/10/023,649

TIME: 13:34:41

Input Set : D:\NEL-0001 Sequence Listing.txt

Output Set: N:\CRF4\01092003\J023649.raw

64	Pro	Asp	Arg	Leu	Gln	Arg	Tyr	Ala	Glu	Arg	Leu	Lys	Lys	Ser	Asp	Ile	
65	90					95					100					105	
67	acc	gac	aag	aac	ata	gcc	tct	aag	gcg	gca	gac	ctg	ctg	gaa	gtc	atg	387
68	Thr	Asp	Lys	Asn	Ile	Ala	Ser	Lys	Ala	Ala	Asp	Leu	Leu	Glu	Val	Met	
69					110					115						120	
71	tca	aca	cca	gac	gca	gag	act	cca	tct	ctg	tgt	atg	cac	aca	gac	gcc	435
72	Ser	Thr	Pro	Asp	Ala	Glu	Thr	Pro	Ser	Leu	Cys	Met	His	Thr	Asp	Ala	
73				125					130					135			
75	acg	tgt	agg	tac	ttt	gga	agt	gta	gca	gta	tac	caa	gat	gtg	tac	gca	483
76	Thr	Cys	Arg	Tyr	Phe	Gly	Ser	Val	Ala	Val	Tyr	Gln	Asp	Val	Tyr	Ala	
77		140						145				150					
79	gtc	cat	gca	ccg	aca	tca	atc	tac	cac	cag	gcg	ctt	aaa	gga	gtt	agg	531
80	Val	His	Ala	Pro	Thr	Ser	Ile	Tyr	His	Gln	Ala	Leu	Lys	Gly	Val	Arg	
81		155					160					165					
83	aca	att	tac	tgg	ata	ggc	ttt	gac	acg	acc	cct	ttt	atg	tac	aaa	aac	579
84	Thr	Ile	Tyr	Trp	Ile	Gly	Phe	Asp	Thr	Thr	Pro	Phe	Met	Tyr	Lys	Asn	
85	170					175				180						185	
87	atg	gca	ggt	tcc	tac	cct	act	tac	aac	acg	aac	tgg	gct	gac	gag	aga	627
88	Met	Ala	Gly	Ser	Tyr	Pro	Thr	Tyr	Asn	Thr	Asn	Trp	Ala	Asp	Glu	Arg	
89				190					195					200			
91	gta	ttg	gaa	gca	cgt	aac	att	ggc	ctc	ggt	aac	tca	gat	ctt	cag	gag	675
92	Val	Leu	Glu	Ala	Arg	Asn	Ile	Gly	Leu	Gly	Asn	Ser	Asp	Leu	Gln	Glu	
93			205					210					215				
95	agc	agg	ctt	gga	aaa	ctc	tca	atc	ctt	agg	aag	aag	agg	ctc	caa	cct	723
96	Ser	Arg	Leu	Gly	Lys	Leu	Ser	Ile	Leu	Arg	Lys	Lys	Arg	Leu	Gln	Pro	
97		220						225					230				
99	act	aat	aag	atc	ata	ttc	tcg	gtt	ggg	tca	aca	atc	tac	aca	gaa	gat	771
100	Thr	Asn	Lys	Ile	Ile	Phe	Ser	Val	Gly	Ser	Thr	Ile	Tyr	Thr	Glu	Asp	
101		235					240					245					
103	aga	tca	ctg	tta	cgt	agc	tgg	cat	ctt	cca	aac	gtg	ttc	cac	ttg	aaa	819
104	Arg	Ser	Leu	Leu	Arg	Ser	Trp	His	Leu	Pro	Asn	Val	Phe	His	Leu	Lys	
105	250				255					260						265	
107	gga	aag	tct	aac	ttc	aca	ggt	aga	tgt	ggg	acc	att	gtc	agc	tgt	gaa	867
108	Gly	Lys	Ser	Asn	Phe	Thr	Gly	Arg	Cys	Gly	Thr	Ile	Val	Ser	Cys	Glu	
109				270					275					280			
111	ggg	tac	gtc	atc	aaa	aag	ata	acg	atc	agc	cca	gga	cta	tac	ggt	aaa	915
112	Gly	Tyr	Val	Ile	Lys	Lys	Ile	Thr	Ile	Ser	Pro	Gly	Leu	Tyr	Gly	Lys	
113			285					290					295				
115	gtt	gag	aac	ttg	gcg	tcc	aca	atg	cat	cgc	gag	ggt	ttc	ttg	agt	tgc	963
116	Val	Glu	Asn	Leu	Ala	Ser	Thr	Met	His	Arg	Glu	Gly	Phe	Leu	Ser	Cys	
117		300					305					310					
119	aaa	gtc	aca	gat	acg	ctg	cgc	ggc	gag	agg	gtt	tct	ttt	gct	gtg	tgt	1011
120	Lys	Val	Thr	Asp	Thr	Leu	Arg	Gly	Glu	Arg	Val	Ser	Phe	Ala	Val	Cys	
121		315				320					325						
123	acg	tat	gta	cca	gcc	aca	ctt	tgc	gat	cag	atg	aca	ggg	att	ctg	gca	1059
124	Thr	Tyr	Val	Pro	Ala	Thr	Leu	Cys	Asp	Gln	Met	Thr	Gly	Ile	Leu	Ala	
125	330				335				340							345	
127	act	gac	gtt	agt	gtg	gat	gac	gca	caa	aaa	cta	ttg	gtt	ggg	ctc	aac	1107
128	Thr	Asp	Val	Ser	Val	Asp	Asp	Ala	Gln	Lys	Leu	Leu	Val	Gly	Leu	Asn	

RAW SEQUENCE LISTING

DATE: 01/09/2003

PATENT APPLICATION: US/10/023,649

TIME: 13:34:41

Input Set : D:\NEL-0001 Sequence Listing.txt

Output Set: N:\CRF4\01092003\J023649.raw

129		350		355		360	
131	caa agg att gtc gtc aat ggt agg acg caa aga aat act aac aca atg						1155
132	Gln Arg Ile Val Val Asn Gly Arg Thr Gln Arg Asn Thr Asn Thr Met						
133		365		370		375	
135	cag aac tat cta tta cca gtg gtc gcc cag gcg ttt tcc agg tgg gcg						1203
136	Gln Asn Tyr Leu Leu Pro Val Val Ala Gln Ala Phe Ser Arg Trp Ala						
137		380		385		390	
139	cgt gaa cat cgt gcc gac ttg gac gac gag aaa gaa cta ggg gtg cgg						1251
140	Arg Glu His Arg Ala Asp Leu Asp Asp Glu Lys Glu Leu Gly Val Arg						
141		395		400		405	
143	gag cgc act ctt act atg ggc tgc tgc tgg gct ttc aag acc cag aaa						1299
144	Glu Arg Thr Leu Thr Met Gly Cys Cys Trp Ala Phe Lys Thr Gln Lys						
145	410		415		420		425
147	atc aca tcc atc tac aag aag cct ggt acg caa aca att aag aaa gta						1347
148	Ile Thr Ser Ile Tyr Lys Lys Pro Gly Thr Gln Thr Ile Lys Lys Val						
149		430		435		440	
151	cct gcc gtc ttt gac tca ttt gtg att cca cgc ctt acc agc cac ggg						1395
152	Pro Ala Val Phe Asp Ser Phe Val Ile Pro Arg Leu Thr Ser His Gly						
153		445		450		455	
155	ctc gat atg ggc ttc cgc cgt agg ctc aag ctg ctg ctt gaa cca act						1443
156	Leu Asp Met Gly Phe Arg Arg Arg Leu Lys Leu Leu Leu Glu Pro Thr						
157		460		465		470	
159	gtc aaa ccc gca ccg gct att aca atg gcc gat gtg gag cat ctg cgt						1491
160	Val Lys Pro Ala Pro Ala Ile Thr Met Ala Asp Val Glu His Leu Arg						
161		475		480		485	
163	ggc tta cag caa gaa gct gaa gaa gtg gct gca gcg gaa gag atc aga						1539
164	Gly Leu Gln Gln Glu Ala Glu Glu Val Ala Ala Glu Glu Ile Arg						
165	490		495		500		505
167	gaa gcc ctg cca ccc ttg ctc cct gaa ata gaa aaa gag acc gta gag						1587
168	Glu Ala Leu Pro Pro Leu Leu Pro Glu Ile Glu Lys Glu Thr Val Glu						
169		510		515		520	
171	gca gaa gta gac ctc att atg caa gag gca gga gca ggt agc gtg gag						1635
172	Ala Glu Val Asp Leu Ile Met Gln Glu Ala Gly Ala Gly Ser Val Glu						
173		525		530		535	
175	aca cca cga gga cac atc agg gtg aca agt tac cca ggc gaa gag aag						1683
176	Thr Pro Arg Gly His Ile Arg Val Thr Ser Tyr Pro Gly Glu Glu Lys						
177		540		545		550	
179	att ggg tct tac gct ata ctt tca ccc cag gcg gta ttg aat agt gaa						1731
180	Ile Gly Ser Tyr Ala Ile Leu Ser Pro Gln Ala Val Leu Asn Ser Glu						
181		555		560		565	
183	aaa ctg gcg tgt atc cac cca ttg gcg gaa caa gta ctg gta atg act						1779
184	Lys Leu Ala Cys Ile His Pro Leu Ala Glu Gln Val Leu Val Met Thr						
185	570		575		580		585
187	cac aaa ggt agg gca ggg aga tac aaa gtc gag cca tac cac ggt aag						1827
188	His Lys Gly Arg Ala Gly Arg Tyr Lys Val Glu Pro Tyr His Gly Lys						
189		590		595		600	
191	gtc att gta cca gaa ggg acg gcg gtc cct gtt caa gac ttc cag gca						1875
192	Val Ile Val Pro Glu Gly Thr Ala Val Pro Val Gln Asp Phe Gln Ala						
193		605		610		615	

RAW SEQUENCE LISTING

DATE: 01/09/2003

PATENT APPLICATION: US/10/023,649

TIME: 13:34:41

Input Set : D:\NEL-0001 Sequence Listing.txt

Output Set: N:\CRF4\01092003\J023649.raw

195	ttg	agt	gag	agc	gct	acg	atc	gtt	ttc	aac	gag	agg	gag	ttc	gta	aac	1923
196	Leu	Ser	Glu	Ser	Ala	Thr	Ile	Val	Phe	Asn	Glu	Arg	Glu	Phe	Val	Asn	
197			620					625						630			
199	aga	tac	ctg	cac	cac	atc	gca	atc	aac	gga	gga	gcg	cta	aac	act	gac	1971
200	Arg	Tyr	Leu	His	His	Ile	Ala	Ile	Asn	Gly	Gly	Ala	Leu	Asn	Thr	Asp	
201			635				640						645				
203	gaa	gag	tac	tat	aag	act	gta	aag	act	cag	gac	aca	gac	tca	gaa	tac	2019
204	Glu	Glu	Tyr	Tyr	Lys	Thr	Val	Lys	Thr	Gln	Asp	Thr	Asp	Ser	Glu	Tyr	
205	650					655					660				665		
207	gtc	ttc	gat	att	gac	gca	cga	aag	tgt	gtt	aag	cga	gaa	gac	gca	ggt	2067
208	Val	Phe	Asp	Ile	Asp	Ala	Arg	Lys	Cys	Val	Lys	Arg	Glu	Asp	Ala	Gly	
209				670						675					680		
211	ccc	ttg	tgc	cta	acc	ggt	gat	ctg	gta	gat	cca	cca	ttt	cac	gag	ttt	2115
212	Pro	Leu	Cys	Leu	Thr	Gly	Asp	Leu	Val	Asp	Pro	Pro	Phe	His	Glu	Phe	
213				685						690					695		
215	gcg	tac	gag	agt	ctc	aag	aca	cga	cca	gca	gca	cct	cac	aaa	gtc	cca	2163
216	Ala	Tyr	Glu	Ser	Leu	Lys	Thr	Arg	Pro	Ala	Ala	Pro	His	Lys	Val	Pro	
217			700					705							710		
219	acc	atc	gga	gtc	tat	gga	gtg	cca	ggt	tca	ggt	aaa	tct	gga	atc	atc	2211
220	Thr	Ile	Gly	Val	Tyr	Gly	Val	Pro	Gly	Ser	Gly	Lys	Ser	Gly	Ile	Ile	
221			715				720					725					
223	aaa	agc	gct	gtg	act	aag	aaa	gat	ctg	gtt	gtg	agt	gcg	aag	aag	gaa	2259
224	Lys	Ser	Ala	Val	Thr	Lys	Lys	Asp	Leu	Val	Val	Ser	Ala	Lys	Lys	Glu	
225	730					735					740					745	
227	aac	tgc	gca	gaa	atc	agg	gat	gta	agg	agg	atg	aga	cgt	atg	gat		2307
228	Asn	Cys	Ala	Glu	Ile	Arg	Asp	Val	Arg	Arg	Met	Arg	Arg	Met	Asp		
229				750						755					760		
231	gtt	gct	gct	agg	act	gtc	gat	tca	gtg	ctt	cta	aat	ggg	gtt	aag	cac	2355
232	Val	Ala	Ala	Arg	Thr	Val	Asp	Ser	Val	Leu	Leu	Asn	Gly	Val	Lys	His	
233				765						770					775		
235	ccc	gtt	aac	act	ctg	tac	att	gat	gag	gca	ttt	gcc	tgc	cat	gca	ggg	2403
236	Pro	Val	Asn	Thr	Leu	Tyr	Ile	Asp	Glu	Ala	Phe	Ala	Cys	His	Ala	Gly	
237			780							785					790		
239	acg	ctg	ctg	gca	ctg	att	gcc	atc	gtc	aaa	cct	aag	aaa	gtg	gta	ttg	2451
240	Thr	Leu	Leu	Ala	Leu	Ile	Ala	Ile	Val	Lys	Pro	Lys	Lys	Val	Val	Leu	
241			795				800					805					
243	tgc	ggg	gac	cca	aaa	caa	tgc	ggc	ttc	ttt	aac	atg	atg	tgc	ctg	aaa	2499
244	Cys	Gly	Asp	Pro	Lys	Gln	Cys	Gly	Phe	Phe	Asn	Met	Met	Cys	Leu	Lys	
245	810					815					820				825		
247	gta	cat	ttt	aac	cat	gac	ata	tgc	act	gaa	gtg	tac	cat	aaa	agc	atc	2547
248	Val	His	Phe	Asn	His	Asp	Ile	Cys	Thr	Glu	Val	Tyr	His	Lys	Ser	Ile	
249				830						835					840		
251	tct	agg	agg	tgc	aca	cag	act	gta	acc	gcc	atc	gtc	tcc	acg	ctc	ttc	2595
252	Ser	Arg	Arg	Cys	Thr	Gln	Thr	Val	Thr	Ala	Ile	Val	Ser	Thr	Leu	Phe	
253				845						850					855		
255	tac	gac	aag	cga	atg	aag	acg	gtt	aac	cca	tgt	gct	gat	aaa	atc	atc	2643
256	Tyr	Asp	Lys	Arg	Met	Lys	Thr	Val	Asn	Pro	Cys	Ala	Asp	Lys	Ile	Ile	
257			860							865					870		
259	ata	gat	acc	aca	ggg	acc	aca	aag	ccg	cac	aaa	gat	gat	ctg	att	cta	2691

RAW SEQUENCE LISTING

DATE: 01/09/2003

PATENT APPLICATION: US/10/023,649

TIME: 13:34:41

Input Set : D:\NEL-0001 Sequence Listing.txt

Output Set: N:\CRF4\01092003\J023649.raw

260	Ile	Asp	Thr	Thr	Gly	Thr	Thr	Lys	Pro	His	Lys	Asp	Asp	Leu	Ile	Leu	
261		875						880					885				
263	acc	tgt	ttc	aga	gga	tgg	gtg	aaa	cag	cta	cag	att	gac	tac	aaa	aat	2739
264	Thr	Cys	Phe	Arg	Gly	Trp	Val	Lys	Gln	Leu	Gln	Ile	Asp	Tyr	Lys	Asn	
265	890					895					900					905	
267	cac	gaa	atc	atg	act	gcg	gct	gca	tcg	caa	gga	ctt	acg	cgg	aaa	ggc	2787
268	His	Glu	Ile	Met	Thr	Ala	Ala	Ala	Ser	Gln	Gly	Leu	Thr	Arg	Lys	Gly	
269					910					915					920		
271	gtt	tat	gct	gtc	agg	tac	aaa	gtc	aac	gag	aat	cca	ctc	tac	tcg	cag	2835
272	Val	Tyr	Ala	Val	Arg	Tyr	Lys	Val	Asn	Glu	Asn	Pro	Leu	Tyr	Ser	Gln	
273			925					930					935				
275	act	tct	gag	cac	gtg	aac	gtg	tta	ctt	aca	cgc	aca	gaa	aaa	cgc	att	2883
276	Thr	Ser	Glu	His	Val	Asn	Val	Leu	Leu	Thr	Arg	Thr	Glu	Lys	Arg	Ile	
277			940					945					950				
279	gtc	tgg	aag	acg	cta	gct	ggt	gat	ccc	tgg	ata	aag	aca	ctt	aca	gct	2931
280	Val	Trp	Lys	Thr	Leu	Ala	Gly	Asp	Pro	Trp	Ile	Lys	Thr	Leu	Thr	Ala	
281		955				960							965				
283	aaa	tat	ccc	ggg	gat	ttc	acg	gct	tca	ttg	gac	gac	tgg	cag	cgc	gaa	2979
284	Lys	Tyr	Pro	Gly	Asp	Phe	Thr	Ala	Ser	Leu	Asp	Asp	Trp	Gln	Arg	Glu	
285	970					975				980					985		
287	cac	gac	gcc	att	atg	gca	cgc	gtt	ctt	gat	aag	ccg	cag	aca	gct	gat	3027
288	His	Asp	Ala	Ile	Met	Ala	Arg	Val	Leu	Asp	Lys	Pro	Gln	Thr	Ala	Asp	
289			990					995					1000				
291	gtg	ttc	cag	aat	aag	gtg	aac	gtc	tgc	tgg	gcg	aag	gct	tta	gag		3072
292	Val	Phe	Gln	Asn	Lys	Val	Asn	Val	Cys	Trp	Ala	Lys	Ala	Leu	Glu		
293			1005					1010					1015				
295	cca	gtc	ttg	gcc	acg	gcc	aac	att	gtg	ctg	acg	aga	cag	cag	tgg		3117
296	Pro	Val	Leu	Ala	Thr	Ala	Asn	Ile	Val	Leu	Thr	Arg	Gln	Gln	Trp		
297			1020					1025					1030				
299	gag	acg	ttg	cac	cca	ttc	aag	cat	gac	aga	gcg	tac	tca	cct	gaa		3162
300	Glu	Thr	Leu	His	Pro	Phe	Lys	His	Asp	Arg	Ala	Tyr	Ser	Pro	Glu		
301			1035					1040					1045				
303	atg	gca	ctg	aac	ttc	ttt	tgc	acc	agg	ttc	ttt	gga	gta	gac	ctg		3207
304	Met	Ala	Leu	Asn	Phe	Phe	Cys	Thr	Arg	Phe	Phe	Gly	Val	Asp	Leu		
305			1050					1055					1060				
307	gac	agt	ggg	tta	ttt	tcc	gct	cct	acc	gtc	gca	ctt	act	tac	agg		3252
308	Asp	Ser	Gly	Leu	Phe	Ser	Ala	Pro	Thr	Val	Ala	Leu	Thr	Tyr	Arg		
309			1065					1070					1075				
311	gat	cag	cac	tgg	gat	aac	tcg	cca	ggg	aag	aac	atg	tat	ggg	ctt		3297
312	Asp	Gln	His	Trp	Asp	Asn	Ser	Pro	Gly	Lys	Asn	Met	Tyr	Gly	Leu		
313			1080					1085					1090				
315	aat	aga	gag	gta	gca	aag	gag	ttg	tca	cgg	cga	tat	ccg	tgc	atc		3342
316	Asn	Arg	Glu	Val	Ala	Lys	Glu	Leu	Ser	Arg	Arg	Tyr	Pro	Cys	Ile		
317			1095					1100					1105				
319	aca	aaa	gcg	gtt	gac	aca	ggc	agg	gta	gct	gat	ata	agg	aat	aat		3387
320	Thr	Lys	Ala	Val	Asp	Thr	Gly	Arg	Val	Ala	Asp	Ile	Arg	Asn	Asn		
321			1110					1115					1120				
323	acc	atc	aag	gac	tac	tct	cca	aca	att	aat	gtg	gtt	cca	tta	aat		3432
324	Thr	Ile	Lys	Asp	Tyr	Ser	Pro	Thr	Ile	Asn	Val	Val	Pro	Leu	Asn		

VERIFICATION SUMMARY

DATE: 01/09/2003

PATENT APPLICATION: US/10/023,649

TIME: 13:34:42

Input Set : D:\NEL-0001 Sequence Listing.txt

Output Set: N:\CRF4\01092003\J023649.raw

L:12 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:2675 M:257 W: Feature value mis-spelled or invalid, <221> Name/Key for SEQ ID#:7
L:2681 M:259 W: Allowed number of lines exceeded, <223> Other Information: